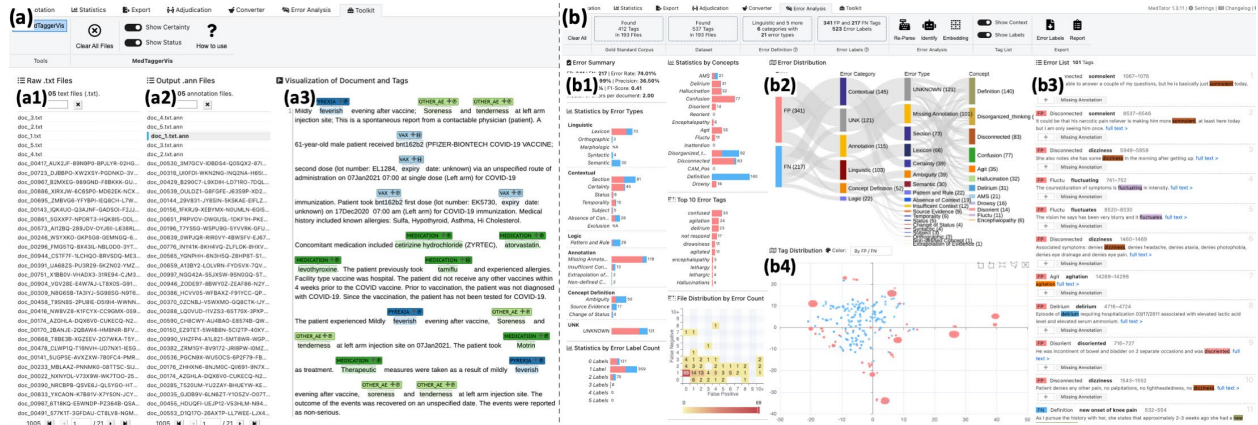


# Visual Exploration of the Results and Errors of Information Extraction

Huan He, PhD\*, Sunyang Fu, PhD\*, Andrew Wen, MS, Liwei Wang, MD, PhD, Sungrim Moon, PhD, Kurt Miller, MS, and Hongfang Liu, PhD

Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, MN, USA

**Introduction.** Information extraction (IE) is a critical step in facilitating the use of healthcare data in clinical decision support and translational research by leveraging natural language processing (NLP) techniques. To evaluate the results of IE systems or algorithms, it is essential to examine the extracted entities and relations with the contextual information. Moreover, as IE errors are not unavoidable, conducting an error analysis is usually required to gain more insights into the reasons for errors. However, it's challenging to analyze the IE errors in its raw format due to the complexity. To address this challenge, we proposed a visual exploration tool to visualize and analyze the IE results and errors based on data visualization techniques.



**Figure 1.** The screenshot of our proposed tool, including (a) the visualization tool for exploring the IE result generated by MedTagger, and (b) the visual analysis of errors between gold standard and the output of the IE system.

**System Design.** To ease the burden of installation, we designed the tool based on a serverless text annotation tool, MedTator<sup>2</sup>, to reuse the in-browser modules for corpus text processing. Users do not need to install any programming language runtime or server to use the tool. As shown in Figure 1, we designed two tabs in MedTator to visualize the IE results (Figure 1a) and errors between the gold standard and IE system outputs (Figure 1b).

To explore the IE results, users can upload raw text files (Figure 1(a1)) and IE result files (Figure 1(a2)) by dragging and dropping them from local disk. Then the user can click on the file name to check the visualized IE system outputs with contextual text (Figure 1(a3)) based on brat visualization<sup>3</sup>.

To analyze the IE errors, users can upload the annotated gold standard, IE system output, and error definition schema into the tool (Figure 1b). Then, the tool can evaluate the errors between the gold standard and IE system outputs. The errors are visualized from multiple aspects, such as error types, error counts, documents, and IE concepts in bar charts (Figure 1(b1)). The number of errors between different aspects is visualized in a Sankey diagram (Figure 1(b2)). By leveraging text embedding and dimension-reduction techniques (t-SNE), the tokens of errors are shown in a scatter plot (Figure 1(b4)) to help users to analyze the semantic similarity of errors. Moreover, users can click on any visual elements in any chart to examine the detailed information of error tokens (Figure 1(b3)). In addition, users can manually label errors using pre-defined error definitions with instantly updated charts (Figure 1(b3)).

The source code and online demo with sample datasets are available at <https://github.com/OHNLP/MedTator>.

**Future Work.** First, we plan to refine the error definition schema for general IE system errors. Users will customize the error definition schema based on their own needs. Secondly, we plan to integrate online text embedding and dimensionality reduction algorithms to improve the performance of error token exploration. Lastly, we plan to conduct usability test with end-users to evaluate the visualization and interactivity designs.

## References

1. Wu T, Ribeiro MT, Heer J, Weld D. Errudite: Scalable, Reproducible, and Testable Error Analysis. In: Proceedings of the 57th Annual Meeting of the ACL; 2019 [cited 2022 Sep 9]. p. 747–63.
2. He H, Fu S, Wang L, Liu S, Wen A, Liu H. MedTator: a serverless annotation tool for corpus development. Bioinformatics, Vol. 38, Issue 6, 15 March 2022, Pages 1776–1778
3. Stenetorp P, Pyysalo S, Topić G, Ohta T, Ananiadou S, Tsujii J. brat: a Web-based Tool for NLP-Assisted Text Annotation. In: Proceedings of the EACL 2012: Demonstrations. Avignon, France: ACL; 2012. p. 102–7.

\* Those authors have contributed equally as co-first authors